R E P O R T   R E S U M E S

ED 020 147   24   SE 004 808

UNDERSTANDING OF CONCEPTS OF PROBABILITY THEORY BY JUNIOR HIGH SCHOOL CHILDREN. FINAL REPORT.

BY- SMOCK, CHARLES D.   BELOVICZ, GRETCHEN
PURDUE UNIV., LAFAYETTE, IND.

REPORT NUMBER BR-6-1295   PUB DATE   FEB 68
GRANT   OEG-3-6-061295-1739
EDRS PRICE   MF-$0.50   HC-$2.52   61P.

DESCRIPTORS- *COGNITIVE DEVELOPMENT, *JUNIOR HIGH SCHOOLS, *LEARNING, *MATHEMATICS, *PROBABILITY, CONCEPT FORMATION, MATHEMATICS EDUCATION, SCIENCE EDUCATION, BRIM SCALE,

INVESTIGATED WERE JUNIOR HIGH SCHOOL STUDENT'S UNDERSTANDINGS OF THE CONCEPTS INTRINSIC TO PROBABILITY THEORY. THE SPECIFIC OBJECTIVES WERE TO EXPLORE (1) THE VARIATION AND CONSISTENCY IN USAGE OF QUANTITATIVE LANGUAGE TERMS, (2) THE ABILITY TO GENERATE POSSIBLE COMBINATIONS AND PERMUTATIONS, (3) THE RECOGNITION AND UTILIZATION OF THE CONCEPT OF INDEPENDENCE OF EVENTS, (4) THE ABILITY TO APPROPRIATELY ASSESS, AND MODIFY, CONTINGENCY RELATIONS IN A COMPLEX SET OF RELATED EVENTS, AND (5) THE EFFECT OF VARIATION IN EVENT-STRUCTURES OF PROBABILISTIC EVENT-SETS ON ESTIMATION OF POPULATION PARAMETERS AND SUBJECTIVE CERTAINTY OF THESE ESTIMATIONS. EXPERIMENTAL PROCEDURES WERE DESIGNED TO OBTAIN DATA ON EACH OF THESE OBJECTIVES. RESULTS INDICATED THAT THE STUDENTS FAILED TO UNDERSTAND THE BASIC IDEA OF PROBABILITY THEORY. RECOMMENDED ARE THAT FURTHER RESEARCH BE UNDERTAKEN ON (1) PARTICULAR CONCEPTS SUCH AS PERMUTATION, CORRELATIONS, AND OTHERS, AND (2) DETERMINING THE EXPERIENCE NECESSARY FOR THE ACQUISITION AND UTILIZATION OF APPROPRIATE QUANTITATIVE LANGUAGE TERMS. (DS)

ED020147

FINAL REPORT
PROJECT NO. 6–1295 PA-24
GRANT NO. OEG–3–6–061295–1739

UNDERSTANDING OF CONCEPTS OF PROBABILITY THEORY BY JUNIOR HIGH-SCHOOL CHILDREN

FEBRUARY 1968

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

OFFICE OF EDUCATION
BUREAU OF RESEARCH

SE 004 808

Final Report

Project No. 6-1295
Grant No. OEG-3-6-06125-1739
9

UNDERSTANDING OF CONCEPTS OF PROBABILITY THEORY BY JUNIOR HIGH-SCHOOL CHILDREN

Charles D. Smock
Gretchen Belovicz

Purdue University
Lafayette, Indiana

February 1968

The research reported herein was performed pursuant to a grant with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

TABLE OF CONTENTS

LIST OF TABLES

ACKNOWLEDGMENTS

Appreciation is extended to all the students and junior high school children who so willingly participated in this study. A special note of appreciation is appropriate for Dr. Gretchen Belovicz who completed a major portion of the study and to Dr. William Belovicz for his kind and helpful assistance. Mrs. Jean Weisbach deserves a special thanks for her patience and skill typing and supervising the final stages of this report.

INTRODUCTION AND BACKGROUND LITERATURE

Problem

The introduction of the "new mathematics" into the educational scheme during recent years has given impetus to many new areas of investigation relevant to educational psychology. The evaluation of its impact on learning and cognitive growth is both complex and controversial. A major question is what prior experience is necessary (stage of development) for the child to have the requisite cognitive skills for understanding the quantitative relationships to be learned. These quantitative relationships can be subsumed under "set theory" which is also the "content" basis of the new mathematical program. The emphasis on set theory lends additional importance to understanding the child's conception and use of concepts of probability. The axioms of probability theory have counterparts in the concepts of set theory, and, in fact, the most classical and rigorous axiomatizations of probability theory are based on set-theoretic terminology and concepts. Therefore, what one learns about the understanding of probability is directly applicable to the understanding of set theoretic concepts.

Statements in the experimental literature imply that young children show some understanding of probability and that Piaget's contention that young children cannot understand probability is questionable (Yost, Siegel, & Andrews, 1962). Other research reports speak of "probability learning" (Messick & Solley, 1957; Stevenson & Zigler, 1958), "strategy" (Jones & Liverant, 1960; Stevenson & Odon, 1964; Yost, Siegel, & Andrews, 1963), and "decision making" (Craig & Myers, 1963; Kass, 1964; Rubin, 1964). In contrast to the implication that understanding is associated with these observations, is the evidence that children of various ages are unable to perform tasks that represent the empirical counter-parts to the axioms and postulates of formal probability theory. If, indeed, the subjects have little or no understanding of the principles basic to probability theory, the assertion that they understand probability as a whole seems highly questionable. Such an assertion may, in part, reflect a confusion between the theoretical model (experimenter's) employed and the underlying cognitive processes

operating. It would seem that a description of the subjects' behaviors based upon the exploration of the complex of skills which underlies the axioms and postulates within probability theory would serve as an aid to understanding reported experimental findings and in the development of appropriate, effective instructional techniques.

Related Research

The experimental study of "probability behavior" began comparatively recently with the introduction of the partial reinforcement technique by Brunswik (1939). The methodology was based, of course, on the Brunswikian assumption that many environmental and event-response contingencies are probabilistic (Brunswik, 1943; Tolman & Brunswik, 1935). The early experiments (with animals) typically involved a binary choice situation in which the level of reinforcement or the ratio of reinforcement was varied between experimental groups. Results focused upon the ability of the animal to discriminate between the various probabilities used, the response levels reached, and the pattern of responses given. Experimentation on probability now has been extended to various species and a number of experimental variations (Jenkins & Stanley, 1950). The term "probability matching" was applied to a widely reported approximation between subjects' responses to a given alternative and the reinforcement level associated with that alternative. This term has been carried over as a descriptive label for human behavior in "probability learning" tasks.

During the 1950's the partial reinforcement paradigm was introduced into developmental research with children and adults. Interest first focused on "probability matching" behavior and success or failure in "maximizing", i.e., consistent responding to the more frequently reinforced alternative. The vast majority of experimental findings indicated that both adults and children tend toward the imput probability level (probability matching rather than maximization).

Several reported findings, however, contradict this body of literature. Messick and Solley (1957) report probability maximization in seven- and eight-year-olds. Uhl (1963) finds that the response level of rats rises above a probability matching level when a large number of trials are used. Edwards (1961) obtains similar results with human subjects, again using a large number of

trials. The pervasiveness of probability matching is not yet fully determined nor is the effect of a number of other variables, such as the effects of reward and punishment, probability preferences, reward and punishment, etc. Further, the 'meaning' of performances characterized by probability matching or maximization is not clear-cut. For example, it has been suggested that what appears to be intelligent maximizing in a child is really the result of the child's lack of experience with probability and consequent failure to perceive the world as an uncertain place (Jones & Liverant, 1960). Piaget (Inhelder & Piaget, 1958) has also postulated a generalized non-differentiation between chance and non-chance events during the preoperational stage.

Another question centers on the degree of understanding with which subjects enter an event-probability learning situation. Although researchers tend to assert that their subjects show some understanding of probability there are several indications of the child's lack of ability to perform tasks closely related to probability. Major deficiencies in related language skills are reported, e.g., inabilities to discriminate indeterminate number concepts (Cohen & Hansel, 1956; Pratt, Hartmann & Mead, 1954), to reason inductively (Welch & Long, 1943), and to perform classifications (Annett, 1959). Numerous studies illustrate the tendency of child and adult to respond to extraneous, interfering events (Smock, 1962), and use response strategies not correlated with the criterial event sequences (Shantz, 1964), e.g., alteration and position preferences. Further, children appear to have limited capacities to discriminate probability levels as well as many gross misconceptions, especially concerning randomness, dependence and independence (Cohen & Hansel, 1956; Gratch, 1959; Inhelder & Piaget, 1958). Studies in our laboratory concerned with the differential role of utility and probability indicate the experimenter's concept of reward maximization many times is at variance with that held by the children (Rubin, 1964; Middleton, 1963).

Numerous studies with adults indicate similar deficiencies. For example, failure to respond to a series as a whole (Estes & Straughan, 1954); failure to learn experimentally even with formal training in appropriate statistical rules (Bennet, Fitts, & Merrill, 1954); the use of non-rational explanation for grouping (Annett, 1959); little or no understanding of correlation (Smedlund, 1963). A study by Brim (1959) is especially

noteworthy. Brim reports that no subjects in a sample of 146 college students were able to correctly combine independent probabilities. Approximately one-third of his subjects made at least one "error" when assigning labels to numerical probabilities, i.e., they placed the label on the wrong side of 50%. At least one investigator has concluded that we should accept the inability of adults to make proper use of information in a betting situation, maintaining that these bets are made in ignorance (Dale, 1962).

The most frequent response in a binary choice situation is that of probability matching, but the evidence reviewed above is sufficient evidence to suggest that many of the "positive" results are a consequence of the particular methodology. The tendency of investigators to "over-interpret" the child's understanding of probability and the data indicating that many subjects have erroneous concepts concerning randomness, utilize extraneous cues and idiosyncratic response strategies, among other problems, indicate a need for research directly concerned with the basic concepts inherent in probability theory.

Objectives

The primary purpose of this study was to investigate junior-high school children's understanding of the concepts intrinsic to the theory of probability. The specific objectives were to explore:

(1) Variation and consistency in usage of quantitative language terms.

(2) Ability to generate possible combinations and permutations, and recognition of all possible outcomes of a particular set of alternative actions of a set of actors.

(3) Recognition and utilization of the concept of independence of events.

(4) Childrens' preferences for low probability events.

(5) Ability to appropriately assess, and modify, contingency relations in a complex set of related events.

(6a) Effect of variation in event-structures of probabilistic event sets on estimation of population parameters and subjective certainty of these estimations.

(6b) Relationship of "need for security" to variation in event-estimation and subjective certainty.

METHOD

Theory

The formal axiomatic treatment of probability theory which is the foundation of the study is outlined in Appendix A. The axioms and postulates presented in the Appendix are taken from Brunk (1960) which is a standard introductory textbook of mathematical probability.

Certain cognitive skills underlying one or more of the axioms have been derived and grouped under three broad categories: Language, Associational Field, the Concept of Independence and Mutual Exclusiveness. A thorough understanding of probability (whether or not construed as formal understanding) must include, as a minimum, an understanding of these factors.

(a) Language

Levels of probability are most exactly expressed in numbers. Some situations demand quantification, but in many situations quantification is either infrequent, impossible, or both; for example, in estimating the probability that a stone lies under a box in the road.

Events in probability theory are made up of elementary events. If event A is made up of more elementary events than is event B, we say it is more probable. With even a simple example, such as the toss of a fair coin, there is a need for distinct but continuous measures of probability. Words or numbers are necessary to describe outcomes. For example: in coin tossing, the universal set has a probability of 1 (heads or tails); head and tail each have a probability of one-half; and the null set is more probable than either heads or tails, each of which is more probable than the null set. One could apply the words "most", "more than", and "least"; terms that may be said to fall on a continuum on which there are an infinite number of possible points.

The number and meaning of such quantification words are necessarily related to the situation and the individual

using them. Fifty dollars may be "many" dollars with reference to the price of a pencil, but "few" dollars with reference to the price of a car. A charitable contribution of ten dollars may be a pittance to a millionaire but a fortune to a pauper. In any case, the child must understand the relative placement of the words on the continuum; i.e., "many" must connote more than "few" within a particular person-situation context.

Ideally, all individuals would have the same set of words to describe quantities (or durations or frequencies). These words would be ordered identically for all individuals. Each word would apply to a range of quantification; moreover, there would be no overlap between words. One might contend that, ideally, the number of words would equal the number of distinct, quantifiable elements.

(b) Associational Field.

The axioms of a probability system are the formal rules that associate the elementary events and the events (combinations of elementary events) with what we call probabilities. Implicit, therefore, is an understanding of elementary events (equally likely, random events) and the rules which govern the combination of these into what we call events (set theory).

An individual can only operate with those elementary events which he can discriminate and recognize as pertinent to the problem under consideration. To the extent that he fails to recognize all possibilities, includes events which are impossible, and/or does not treat the elementary events as being equally likely, he will fail to use and understand probability properly.

Once a set of elementary events has been established it is necessary to form all the combinations of the elementary events. The individual must consider the universal event, the null event, and all combinations between them.

(c) Concept of Independence and Mutual Exclusiveness.

Events are mutually exclusive if they have no element in common. In the case of a die, the occurrence of either a one or a two might be considered as one event and the occurrence of either a five or a six as another event. These events are each made up of two elementary events or elements and are mutually exclusive since either the first event can occur or the second event can occur, but not both on a single trial.

The concept of additive probabilities is associated with mutually exclusive events. Additivity does not hold where mutually exclusive events do not exist, since some elementary event would be counted more than once. Thus, the probability of the event "one" or "two" from a toss of a die would be the sum of the probability of the first and the probability of the second. Generally, in talking about mutual exclusiveness, we are considering a single simultaneous observation. The addition of probabilities where events are not mutually exclusive presents an obvious source of error in the utilization of probability.

Two events are independent if the occurrence or non-occurrence of one event has no bearing on the occurrence or non-occurrence of the other. The outcome of a toss of a coin is not affected by the prior outcomes and we act as if these are independent events. However, if in fact the manner in which the person flips the coin is related to the side which appeared the last time the coin was flipped, the outcomes are dependent.

The concept of independence should not be confused with the change in the estimate of the probability of an event. For example, Cohen and Hansel (1956) describe the behavior of children in a probability estimating situation. The child subjects are shown an urn and told it has beads of two colors. The beads are drawn one at a time and after each draw the child is to predict the color of the next bead drawn. A sequence of nine yellow beads had been drawn and a frequent prediction was "yellow". The children's explanations of their choices were on the basis of the predominance of yellow beads previously drawn. Cohen and Hansel explain this phenomenon as a lack of understanding of the concept of independence. An equally good explanation (Bayesian) would be that the subjects assume a greater proportion of yellow beads in the urn.

Experimental Tasks

The methodology employed in this study was aimed at investigating the considerations described in the previous section. Many of the concepts are highly inter-related and a design utilized that permitted investigation of these inter-relationships. The study uses a relatively large number of subjects but it is planned to provide as complete a set of data as is possible for some children. Thus, it relies on group testing in its initial stages, with a second stage devoted to individual testing.

The choice of the subject population was based on the research of Piaget and his associates. The work of Piaget strongly suggests that the ability to perform logical operations, such as are involved in the use of probability, appears during pre-adolescence and early adolescence (Inhelder & Piaget, 1958). Twelve- and thirteen-year-old children should exhibit some, but not all, of the relevant competencies and should exhibit these to varying degrees. Further, they represent a group in which verbal skills are well developed, which increased the utility of inquiry.

1. Quantitative Language Usage

    1a. Language - The assignment of quantitative terms by the subject to given lists of fractions, percentages, absolute numbers (small range), and absolute numbers (large range). Instruction were general so as to minimize structuring of word associations. Three types of quantity and two range levels were used for comparative purposes.

    1b. Language - The assignment of absolute numbers, percentages, and fractions to quantitative words (quantifiers) by the subject. Lengths of lists and instructions will be similar to those used in Task 1a. The task provided an index of associations and were used for purposes of comparison with responses to Task 1a.

    1c. Language - Matching quantifiers with absolute numbers, percentages, and fractions. Lengths of lists was similar to those used in Task 1a. Instructions were to match each element once only. The task provided an index to order tendencies and was used in comparisons between responses in structured and unstructured fields of choice. The internal consistency of Ss's behavior can be studied by comparison of Tasks 1a, 1b, and 1c.

For Task 1a, columns (a) through (d) are examples of the lists shown the subjects. Column (e) is an example of the lists for Task 1b, while Task 1c consisted of a Task 1a and a Task 1c list presented simultaneously.

Table 1

Sample of Language Usage Tasks

(a)	(b)	(c)	(d)	(e)
1/4	20%	1	1	Hardly any
1/3	35%	5	117	Few
1/2	53%	17	3312	Several
2/3	65%	50	10,151	
3/4	90%	100	100,000	Some
				Many
				A lot
				Almost all

2. Combination and Permutations

2a. Operations - The formation of permutations in three element groups. The task provides an index to ability to form all arrangements of the elements.

2b. Operations - The formation of all possible pairs in a three element group. The task provides an index to the use of combinations vs. the use of permutations since it requires only one response, i.e., the subject may choose to respond with permutations or with combinations.

2c. Operations - The formation of all possible pairings in a three element group different from the formation given in response to Task 2b. This task measures the ability of those Ss using one method of grouping in Task 2b to recognize and use the alternative grouping.

Table 2

Sample of Combination and Permutation Tasks
(Tasks 2a, 2b and 2c)

For tasks 2a, 2b and 2c, Ss were given three symbols, such as a circle, triangle, and square, and are asked to perform the indicated operations.

Directions:

(Task 2a):	"Arrange the three figures in all the ways you can."
(Task 2b):	"Pair up the three figures in all the ways you can."
(Task 2c):	"Find a way to pair up the three figures which is different from the way you used in question 2b."

2d. Operations - Recognition of all possible outcomes in an imaginary game involving three players with four alternative actions available to each player (i.e., no action, Action A, Action B, Action C). The task measures the subject's ability to recognize not only combinations and permutations (as in Tasks 2a, 2b and 2c) but also to enumerate all events in the probability that are composed of elementary events from the element space (A,B,C).

Table 3

Sample of Ability for the Enumeration of Possible Events
(Task 2d)

Directions:

"Ten children are entered in a contest. At the end of the contest, the judge will announce the names of those who win the first, second, and third prizes. Three of the contestants, John, Bill, and Hank, are from your school and you are most interested in how they do in the contest."

"Show all the things which could happen to the three boys. Use the numbers 1, 2, and 3 to mean first, second, and third prize and a 0 to mean no prize. For example, if John won, Bill came in third, and Hank did not place, you would put down a 1 under John's name, a 3 under Bill's name, and a 0 under Hank's name."

3. Concept of Independence and Mutual Exclusiveness

3a. Independence - The task to measure ability to recognize the independence of possible outcome consisted of estimations of mathematical probabilities in problems involving: 1) a coin toss, 2) a die throw, 3) a car race, and 4) a raffle. Both a multiple choice and fill in format were used.

3b. Independence - The estimation of probabilities in problems analogous to those used in Task 3a, using verbal solutions was used to measure the recognition of independence of possible outcomes without requiring correct mathematical solutions.

3c. Independence - This task measured the recognition of independence as expressed through the subject's choice in a simulated betting situation was determined by S's comparison of likelihoods in problems analogous to those used in Task 3a.

Table 4

Sample of Independence of Events Task
(Tasks 3a, 3b, 3c)

(3a) If you flip two coins, heads will come up on both

- a) half of the time
- b) always
- c) never
- d) one-fourth of the time
- e) other
- f) don't know

(3b) If you flip two coins, tails will come up on both what fraction of the time?

- a) 1/2
- b) 1/3
- c) 1/4
- d) 2/3
- e) other
- f) don't know

(3c) If you flip two coins you will get two heads

- a) more often than a head and one tail
- b) less often than one head and one tail
- c) the same number of times as one head and one tail
- d) don't know

4. Low Probability Events.

Subjects were asked to express preference for events of equal expected value when one event includes an element of low probability. A series of choice situations was couched in simple concrete terms (such as wagers). The task measures the tendency of the subject to be adversely influenced by a low probability element, despite equal expected outcomes for the two events.

Table 5

Sample of Low Probability Event Preferences Tasks

Directions: S is asked to choose the team be believes has the best chance of winning.

Team 1 Members	Team 1 Chances of winning points	Team 2 Members	Team 2 Chances of winning points
1	9/10	1	9/10
2	6/10	2	6/10
3	6/10	3	5/10
4	6/10	4	4/10
5	1/10	5	4/10

5. Correlation.

The tasks dealing with correlation replicate a methodology used by Piaget with a similar age group (Inhelder & Piaget, 1958). In a set of 40 pictures, children are portrayed with brown hair and brown eyes, or blond hair and blue eyes, or brown hair and blue eyes, or, fourth, with blond hair and brown eyes.

5a. Correlation - Estimation of the strength of relationship existent in given subsets. S verbally

described the strength of the relationship.

5b. Correlation - Estimation of the comparative strengths of relationships existent in subsets presented two at a time. S made verbal comparisons.

5c. Correlation - Ability to strengthen a relationship existent in a given subset through the addition or subtraction of instances. S was asked to manipulate the stimulus pictures in order to produce requested changes.

6. Information.

The random order of presentation, regular order of presentation, full information, summarized information, single draw, multiple draw, and length of run was studied. Full information is defined as constant access to the visual display of all prior outcomes and observation of each draw as it is made. Summarized information is defined as access to only the numerical total of outcomes in a series, with no access to a visual display to prior outcomes and no observation of actual draws. Short runs are terminated at an arbitrary point. Long runs were terminated when response stability is achieved. The tasks all involve binary-choice situations.

6a. Information - The estimation of population ratios before a draw is made and after subsequent draws, the prediction of outcome of each draw.

6b. Information - The explanation of population estimates and predicted outcomes made by an imaginary person in tasks identical to those used in sampled portions of Task 1. The task investigates the range and type of hypotheses recognized by the subjects when personal involvement is minimal.

Subjects

Thirty-six college (10 F, 26 M), and 35 junior-high school students served as subjects in the experiments. Subjects in both samples were selected randomly from populations of approximately 150 students. Relevant descriptive data are presented in Appendix B.

Procedure

Preliminary investigation indicated a necessity for fuller exploration of the role of event-structure and other information on estimation of population ratios. Therefore, this report consists of the results of three separate studies designed to explore the objectives outlined above. The specific procedures for each of the three studies is described in the appropriate section of the Results. The first study examined the role of event-structure and rationality (i.e., information utilization) of college students' in probabilistic situations (Objective 6a and 6b). The second study was concerned with achieving objectives one through four, whereas the third examined the understanding of correlations (Objective 5) and information utilization (Objective 6) in junior high school subjects.

RESULTS AND DISCUSSION

I. Information: Event Structure and Population Estimation

Preliminary investigation of the suitability of the experimental tasks indicated a thorough analysis of the role of event-structure on estimation of population proportion, certainty of those estimations and the S's explanations (rational or not) (Objectives 6a and 6b) for their estimation would be necessary in order to adequately administer the task to junior high school students. This first study*, therefore, was specifically designed to determine: (1) the effect of different presentations of quantitative data on estimation of population parameters and the subjective certainty of those estimations; (2) the extent rationality; and (3) the relationship of responses in the probabilistic situations presented and "need for security".

Subjects were 36 college students (26 male and 10 female) of average ability level varying mathematical backgrounds.

Three experimental tasks (A,B,C) were designed to study the effect of: a) event frequency or sample proportion, b) event structure (order), c) access to summary information relevant to event frequency and order, d) sample size, and e) display pattern on estimation and certainty responses. A fourth task (D) was a test designed by Brim (1955), to study the effect of "need for certainty" on behavior in the probabilistic situations.

Task A consisted of seven data arrays of 30 elements each. Four of the arrays contained 50% red and 50% blue elements and three arrays had a sample proportion of 33-67. Order of the elements in five arrays was randomized whereas sequential dependencies (i.e., alternation) existed in the remaining two arrays.

Fifteen arrays of 150 elements each were presented in Task B. Sample proportions of 50-50 and 33-67 were

---

*Complete data analysis and conclusion from this study may be found in: Belovicz, Gretchen J. Subjective Beliefs and Probability Behavior. Doctoral Dissertation, Purdue University, 1967.

again used and order was varied as in Task A. Task C differed from Task A (and B) in that summary information was given the subject along with the displays. This information consisted of a count of the red and blue elements (totals), event contingencies in the form of a numerical contingency table, and event contingencies in the form of sentences (verbal contingency table).

In Task D (Brim Scale, CF Brim, 1959) the subjects were presented with 32 statements dealing with the probability of certain "action outcomes". The statements varied with respect to content and actual probability.

All data displays were presented to Ss in "booklet" form. Data elements (red and blue circles) were drawn on 9" x 11" sheets of white paper, which were covered with clear plastic. These displays were inserted face-up on the left side of the booklet. The scale for Certainty Judgments were printed on the page facing the displays.

Elements in Task A displays were placed in perpendicular double-columns with elements of one color placed slightly to the left of the other.

Element displays for Task B consisted of five perpendicular double columns except in the six displays involving letters of single columns. Displays used in Task C were identical to those used in Task A, except for the summary information placed at the bottom of the opposite page (beneath the Certainty Scales). The color of initial elements and the placement of red and blue elements in the right or left column varied across displays.

Response Measures

The following response measures were obtained on Tasks A, B and C: 1) estimation of the proportion of red and blue elements in the population from which the sample (display elements) was drawn, 2) the subjective certainty with which these estimates were made, (five-point scale), 3) decision time (latencies) for the estimation of population proportions, 4) decision time (latencies) for the "certainty" estimates, and 5) explanations of Ss relevant to rules and procedures they used for both estimations and certainty judgements.

The Brim test requires Ss to estimate the objective probability of each of the 32 events described and to

indicate the certainty of these values on a five-point scale. Thus, the dependent variables in this task were analogous to variables (1) and (2) in Tasks A, B, and C.

Procedure

Subjects were tested individually in one hour sessions. All but six subjects were tested by one of the experimenters (GB). A small room equipped with table, chairs, and tape recorder, was used for data collection.

The nature of the experiment was described briefly to the subject and the hypothetical technique for sample selection was explained to him. S was asked to look at each display in Task A and to tell the experimenter (E) "about how many" red and blue elements he thought were in the total population of 1000 elements. Following his response, S was asked to indicate the certainty of his judgment on the five point certainty scale.

The same procedure was used for Task B, but S was forewarned that the displays contained 150 rather than 30 elements. In Task C, S's attention was called to the summary information and the contingency tables were thoroughly explained to him. Explanation was continued until E was sure that S understood the significance of the information.

After responses had been obtained to all displays an Inquiry period was conducted and recorded on tape. During the Inquiry period S was asked to complete the Brim scale. Finally, personal data relating to age, class placement in the University, academic field of concentration, and background in mathematics were obtained.

Qualitative Results

Data obtained from the Inquiry were analyzed to determine: (1) strategies used in the estimation of population proportions on Tasks A and B, (2) the use of information sources in Task C, and (3) justification and reasons for variations in certainty judgments.

Strategies

Strategies used in the estimation of population proportions were described by subjects with varying degrees

of clarity, elaboration, sophistication. The strategies could be classified as follows:

I. Counting. Ss counted all or part of the display elements. On small samples (Task A) Ss generally counted elements of one color only. Few Ss counted elements on alternating displays. No S counted all elements in Task B displays.

II. Visual Estimation. Ss relied on over-all impression and generally worked rapidly. Visual estimation was differentiated from other strategies involving visual observation by the cursoriness with which it was carried out. In some cases explanations were placed in this category because S was unable to explain his methodology more explicitly.

> Subject 1: ". . . . I approximately tried to size up with my eye the colors, according to what the bulk of the color, like if I thought there was about two times more blue as there was red I used like 700 and 300 . . . . like if there is more blue than there was red I used that as a kind of a proportion."

III. Comparison of Color Blocks. Ss scanned the display for "groups" or "blocks" of one color, which they then compared to "blocks" of the other color. Seldom was a "block" defined as having less than three of four elements. This strategy had much in common with Visual Estimation; however, these Ss proceeded somewhat more systematically.

> Subject 16: ". . . . I'd go along and I'd see a group of red ones, say 4 or 5. I'd see if there was a counteracting group of blue ones along there . . . ."

It should be noted that Ss believed a "block" of one color, in a display with proportions of 50-50 should be matched with an equally long "block" of the other color. Thus, they failed to consider the possibility that a "block" of four red elements might be "matched" by two "blocks" of two blue elements each.

IV. Most Striking Color. Ss based estimates on their first feeling of color dominance. The methodology

was impressionistic and allowed the S to work very rapidly. This strategy had much in common with Visual Estimation, but was more inaccurate.

> Subject 10: ". . . . I just decided which color struck me first when I first looked at it. If there was a perponderance of blue or red then however much I noticed it more than the other."

V. Column-by Column Proportions. Ss chose one column, counted elements or visually estimated a proportion from that column. Ss sometimes compared the results to proportions estimated from one or more of the remaining columns. In some cases choice of the first column was mechanical, i.e., the S proceeded from left to right. In other cases, choice was based upon representativeness, i.e., the most or the least representative column was chosen. In those cases where Ss proceeded mechanically from left to right the strategy closely resembled Visual Estimation and could be differentiated only by the more systematic nature of the approach. Differences between this strategy and Visual Estimation were more pronounced when Ss sampled the display by choosing one or more columns on the basis of representativeness. This strategy, of course, could only be used on Task B displays. Further, by its very nature it had to involve some other strategy, i.e., the elements within each column had to be estimated or counted in some fashion.

VI. Patterning. Explanations placed in this category were somewhat crude with considerable variations but Ss explicitly referred to the sequence characteristics of the displays.

The frequency with which each strategy was used in Tasks A and B, whether alone or in combination with some other strategy, is given in Table 6.

Table 6

Frequency of Each Strategy on Tasks A and B

Strategy	Frequency	
	Task A (N=38)	Task B[a] (N=52)
Counting (N=24)	20	4
Visual Estimation (N=32)	10	22
Comparison of Color Blocks (N=13)	3	10
Most Striking Color (N=5)	2	3
Column-by-Column Proportions (N=12)	0[b]	12
Patterning (N=4)	3	1

[a]Twelve of the Task B cases represent a combination of Column-by-Column Proportions and Counting or Visual Estimation.

[b]Inapplicable on Task A displays.

Utilization of Information

Displays in Task C provided subjects with four potential sources of information: (1) the display itself, (2) the numerical totals, (3) the numerical contingency table, and (4) the verbal contingency table. Table 7 presents the frequency with which information sources were used alone or in combination.

Table 7

Single and Combinatorial Use of Information Sources

Information Source	Frequency (N=36)
Display	1
Totals	7
Table, Numerical	1
Table, Verbal	1
Display + Totals	7
Display + Numerical Table	3
Display + Verbal Table	1
Display + Totals + Table (Num.)	8
Display + Totals + Table (Verb.)	1
Display + Tables	0
Totals + Table (Num.)	5
Totals + Table (Verb.)	0
Totals + Tables	1

It is apparent from the table above that three-fourths of the subjects used combinations of the information sources. The numerical totals were used most frequently whereas the verbal contingency table was the least popular information source.

The contingency tables were rejected as a source of information by about fifty per cent of the subjects. Ss expressed little confidence in justifying the rejection and their reasoning was usually fallacious.

There was no statistical relationship between choice of an information source(s) on Task C and type of strategy on Tasks A and B.

Variations in Level of Certainty

Subjects attributed variations in certainty to sample size, alternation, sample proportions, display pattern,

and the presence of information. The total frequency with which display characteristics were mentioned is given in Table 8.

Table 8

Explanations for Variation in Certainty Judgments

Display Characteristics	Frequency (N=98)
Alternation	28
Sample Size	27
Information	23
Proportion	12
Display Pattern	8

About 50 per cent of the Ss gave more than one reason for a particular certainty estimate. The following display characteristics being most frequently associated with an increase in certainty: (1) small sample size, (2) alternation, (3) 50-50 proportions, and (4) presence of the quantitative information.

Ss manifested a strong tendency to discuss certainty judgments in terms of what they had done and not what they had seen in the displays. The concentration on the mental activity and strategies is consistent with the Ss emphasis on the value of obtaining accurate sample proportion estimates. For example, higher levels of subjective certainty was significantly associated with Ss belief they had achieved greater accuracy in estimates of sample proportions.

Quantitative Results

(1) The majority of Ss were able to discriminate between population proportions of 50-50 and 33-67. However, some overlapping of responses did exist and a few Ss reported the same proportion in 50-50 displays as others reported in 33-67 displays. Response accuracy was

greatest on 50-50 alternating displays (98% correct responses). Element alternation did not increase accuracy of estimations.

(2) Subjects tended to overestimate proportions about .50 and underestimate proportions below .50, a finding consistent with data from prior research.

(3) Large sample size was associated with decreased certainty, primarily as a result of the Ss reliance on accurate sample proportion estimation together with a decrease in accuracy associated with large samples.

(4) Randomization of display elements also were associated with decreased certainty.

(5) There was a strong "preference" for proportions of 50-50 under all conditions. Responses to the Brim scale also tended to cluster around .50 even though the actual proportions were distributed rectangularily.

Certainty on the experimental tasks tended to increase as proportion estimates approached .50 ($r = -.13$, $df = 36$, $p < .001$) when individual data was analyzed. This was contrary to Brim's hypothesis that certainty increases as estimates approach the extremes of 0 and 100%. During the Inquiry session some subjects evidenced a belief that proportions of 50-50 were "natural" or "better".

(6) Associations between latencies, accuracy, and certainty were inconsistent. Latency measures on all the tasks appeared to reflect complex interaction effects of the independent variables, practice and motivation.

(7) Reliability of response was generally good. Comparisons between responses to similar displays yielded significant correlation when the first displays on Task A, B, and C were involved. Unusual responses, which reduced individual response consistency, was elicited on tasks representing different display variables.

(8) Subjects tended to respond to both the experimental tasks and the "desire for certainty" measure at a given level of certainty ($r = .48$, $df = 36$, $p < .01$). This level of certainty appears related to the postulated (Brim, 1959) "need for security".

Conclusions

It may be concluded from this first study that responses in these "probabilistic situations" could not be taken at face value, that is, labeled *per se* as "good" or "bad", "appropriate" or "inappropriate". The rationality of the response could only be evaluated through investigation of the decision process underlying the responses and a "subjective beliefs" of the subject. In many cases, such an analysis of responses indicated *Ss* "predictions" derived from erroneous subjective beliefs.

The college students participating in this study gave little evidence of understanding probabilistic concepts. The findings contradict assumptions implicit and/or explicit, of some experimenters concerning rationality of the behavior of children in probability learning" situations (e.g., Yost, Sigel and Andrews, 1962).

Study Two: Children's Usage of Quantitative Language Terms and Selected Concepts of Probability Theory.

The purpose of this study was to determine children's usage of quantitative language terms and understanding of selected concepts of probability (permutations and combinations, independence, response to low probability events). As indicated in the introduction, there is little evidence that children understand the concepts mentioned above or manifest consistent language usage with respect to quantitative expressions. The investigation is preliminary in nature; neither the review of the literature nor the prior study (Belowicz, 1967) yielded any basis for a more systematic investigation. The specific questions to which we addressed ourselves were: 1) What types of quantitative language terms do children have available, and to what extent are these terms used with inter- and intra-individual consistency? 2) To what extent do junior high school children understand the processes underlying permutations and combinations? 3) To what extent do junior high school children understand the concept of independence of probabilistic events? 4) To what extent do children utilize low probability events?

Procedure

Several items representing each of the selected concepts of probability theory were constructed and, following pre-testing, a final Probability Inventory was constructed (See Methods, pp. 8-14 for sample items). The Inventory was administered, in a group situation, to 63 children (35 boys and 25 girls) in two eighth grade mathematics classes of a local junior high school. The Inventory consisted of four parts (a set of items relevant to each of the questions above) assembled in two orders counter-balancing the types of items.

Results and Discussion

Part I: Utilization of Quantitative Language Terms

The nature of the results, the data, and the results obtained precluded precise kinds of quantitative analysis. Therefore, the results, for the most part, are presented in descriptive and narrative form.

In Part I of the inventory, Ss were asked to supply language term equivalents for sets of fractions, percentages, or absolute numbers and then to match a set of language terms sets of these three forms of numerical expression. Each set of numerical terms ranged from "low" to "high" with the range varying across sets or "items".

The first finding of interest was the great variety of language terms associated with each level of numerical terms within an item set. The number of terms ranged from 19, at level three, to 31 at two of the other levels. Inspection of the list (Appendix C) indicates that Ss utilization of quantitative language terms varies from quite immature and imprecise (e.g., little, little bit) to precise expression of the relational aspect of the number system within a set (e.g., fewest). Level three of the range was particularly interesting. Many Ss were able to give fairly precise numerical expression to the characteristics to the numerical expressions at that level. For example, several Ss would report this particular subitem as representing "about 1/2", or average amount. Although the quantitative range of the sets (i.e., the "item"; fractions, percentages, or absolute number) varied by the order of at least three, analysis indicated variation in range had no appreciable effect on Ss language behavior.

A crude estimate of individual consistency in utilization of language terms across the different kinds of sub-items (i.e., fractions, percentages, and absolute numbers) was obtained. Five levels of consistency were defined: Level 5 (very consistent) was defined as those responses that were identical for particular points of the range independent of specific item or type of item presented. Level 4 was defined as being of the same level of quantitative language usage and consisting of synonyms for identical points in the range across items and types of items. Levels 3, 2, and 1 (lowest consistency) were defined in terms of increasing inconsistency in the use of language terms for identical points in the sub-set of items and types of items. For example, if the lowest point in the range consisted of one/tenth (i.e., 1/10 or 10%), and 10, with 100 being the highest for the absolute number set, a S who responded with "almost none", "some", and "good size" was scored at Level 1. Table 9 presents the results of this analysis. Only 50 percent of the Ss achieved a Level 4 or 5 of consistency. Analysis of the matching items yielded similar indications of a lack of intra-individual consistency for those children obtaining low consistency scores on the "free response" items.

Table 9

Consistency Score Across All Types of Sets of Numerical Items

		N	%
Inconsistent	1.	9	13
	2.	13	21
	3.	17	27
	4.	21	33
Consistent	5.	3	5

We conclude that junior high school students vary significantly in the adequacy of appropriate usage of

language terms and manifest considerable variability in the appropriate use of relational and comparative terms. There also was relatively little degree of intra-individual consistency in the use of these terms across different types of numerical expressions. About 50% of the children were fairly consistent, though they were not yet utilizing the same or exact quantitative expressions. When Ss were given language expressions and asked to express these in numerical terms (either fractions, percentages, or absolute numbers) the results were the same. It seems clear that the children's understanding of numerical terms was less than might be expected. Whether lack of understanding of the number system or inadequate availability of quantitative language terms is the basic contributor to the inconsistency and inaccuracy cannot be answered by these data.

Part II

Part Two of the Probability Inventory consisted of a number of items requiring the children to respond to possible arrangements (permutations) and combinations of letters, symbols and numbers (see Methods, pp. 10-11).

Table 10presents the mean and standard deviation of correct, incorrect, and repetitious permutations and combinations for each item of Part II.

Table 10

Mean and Standard Deviation of Correct, Incorrect and Repetitious Responses for Each Item of Part II

Items			Type of Response		
			Correct	Repetitious	Wrong
1.	L - P*	mean	5.60	.27	.61
		st. dev.	1.21	.42	3.10
2.	L - C	mean	5.50	.13	.21
		st. dev.	1.21	.33	.67
3.	L - C	mean	.21	.82	.87
		st. dev.	.62	2.20	1.41
4.	S - P	mean	5.40	.30	.57
		st. dev.	1.18	.47	.26
5.	S - C	mean	5.50	.12	.19
		st. dev.	1.22	.34	.67
6.	S - C	mean	.11	1.00	.98
		st. dev.	.40	1.64	1.40
7.	N - P	mean	5.61	.16	.70
		st. dev.	1.20	.41	3.20
8.	N - C	mean	5.54	.06	.39
		st. dev.	1.12	.08	.89
9.	N - C	mean	.33	.88	.65
		st. dev.	.73	2.00	1.20

*Key: L - Letters, S - Symbols, N - Numbers
P - Permutations, C - Combinations

Analysis of these data indicated no statistically significant variation could be attributed to IQ level, mathematical background or sex of S, nor to order or presentation of the items.

The percent of correct, incorrect and repetitious responses to groups of similar items of Part II is presented in Table 11.

Table 11

Percent of Responses Correct, Incorrect and Repetitious for Groups of Similar Items of Part II

Type of Item	Correct (%)	Incorrect (%)	Repetitious (%)
Three Elements - P	92.5	4.15	10.5
Two Elements - C	93.0	1.7	4.4
Three Elements - C	3.6	15.1	13.9
Letters 1,2,3	63.0	6.5	9.4
Symbols 4,5,6	60.8	8.0	9.7
Numbers 7,8,9	64.8	6.2	9.7

It will be noted that there is relatively little variation attributable to the type of element (i.e., letter, symbol, or numbers), although there was a tendency for more incorrect responses to be associated with items involving symbols. Permutation of three elements appeared to be relatively easy for this age group, however more incorrect and repetitious responses were elicited. The combination of two elements, as expected, was significantly easier than tasks involving three elements. It was clear from inquiry data that few of the Ss actually understood the basis concept of combination and were operating on a trial and error basis. The low percentage of correct responses, the many incorrect and repetitious responses reflect this lack of understanding.

Table 12

Percent Correct, Incorrect and Repetitious Permutation for Each Item of Part II

Items	Type of Response		
	Correct (%)	Incorrect (%)	Repetitious (%)
1. L - P*	93.0	4.5	10.3
2. L - C	92.1	2.1	3.4
3. L - C	3.4	13.5	14.5
4. S - P	90.5	5.3	9.5
5. S - C	92.5	2.1	3.2
6. S - C	1.8	16.7	16.4
7. N - P	94.0	2.6	11.5
8. N - C	95.0	1.0	6.6
9. N - C	5.5	14.8	10.8

*Key: L - Letters, S - Symbols, N - Numbers,
P - Permutations, C - Combinations

These results are further confirmed by the findings from the final section of Part II. The final two items involved the recognition and generation of all possible combinations of events. The items depicted two situations; one involving two girls in a baking contest and the second of three boys entered in an athletic contest. Ss were asked to indicate, in each case, all the possible outcomes that could occur to the children in the particular contest. Table 13 presents the percentage of correct, incorrect and repetitious responses for each item in each IQ level group.

Table 13

Percent Responses Correct, Incorrect and Repetitious by IQ Level and Items 1 and 2 (Part III)

	Correct (%)		Incorrect (%)		Repetitious (%)	
IQ Level Item	1	2	1	2	1	2
80-104(N=20)	69.2	25.2	5.8	16.2	12.5	19.9
105-114(N=21)	78.6	30.7	2.1	10.3	26.1	8.0
115+ (N=20)	90.3	39.7	2.0	1.9	8.6	4.4

The number of elements to be combined (i.e., 2 or 3 and yielding respectively 6 or 24 combinations) and IQ level both proved to be highly significant factors in levels of performance. The high level of repetitions on item 2 for both the average and low-average intelligence level groups indicates clearly their lack of understanding of combinatorial principles. At the same time, the relatively low percentage of the potential combinations generated by the higher IQ groups (39.7%) indicates these Ss also do not have complete understanding of the necessary principles of combinations.

Part III

The third part of the Probability Inventory consisted of 9 items (5 in quantitative and 4 in verbal terms) concerned with the children's understanding of the concept of the independence of probabilities. As Table 14 indicates verbal items were slightly easier than those presented in quantitative terms.

Table 14

Percent of Correct, Incorrect and Don't Know Responses on Independence of Probability Items (Part III)

	Correct (%)	Incorrect (%)	Don't Know (%)
Quantitative	26.2	48.4	25.4
Verbal	32.6	38.7	28.7

There were no highly significant differences, though a trend of the higher IQ children performing better than those in the lower IQ groups. Further analysis indicated a strong tendency for Ss to prefer "1/2"; an answer more often than not was incorrect. It should be pointed out that 10 Ss (not in the same IQ level group) contributed a large number of the correct responses. In addition, comparison of the identical items stated in different forms (i.e., quantitative or verbal) indicated all Ss performance (correct or incorrect) was particularly consistent on these items.

Part IV

The final section of the inventory presented the S with 5 items designed to determine preference or avoidance of low probability events. Specifically, in each case the child was to indicate which of two teams were most likely to win a contest when the individual members of each team had differing probabilities of winning points (see Methods, pp. 11). In each case, one of the teams would have individuals with high variability of performance possibilities (for example, Member 1 might have 9 chances out of 10 of winning whereas Member 5 would only have 1 chance out of 10 of winning) while the members of the other team were more equally distributed in terms of "skill" and possibility of winning. The probability of the team winning was exactly the same for both teams in all items.

Analysis of these results indicated no significant effects attributable to order of presentation, sex of S, IQ level, or age of Ss. There was a slight tendency for those "teams" to be preferred which had low variability among the individual probabilities, but the percentages of choice among the two items of each set was generally close to the 50% mark. Post-test inquiry indicated Ss reasons for making choices were unrealistic and illogical. Only a few Ss (12%) consistently indicated that both teams would win (i.e., had an equal chance of winning).

The results obtained from the Probability Inventory may be summarized quite briefly. While it was not feasible in most cases to use more than descriptive statistics, it is clear that the order of presentation, sex of S, and IQ level (except in one or two instances) was not a significant determinant of performance level for junior high students on this inventory. The inventory yielded indications of a wide variety of available quantitative language

terms but in over 50% of the cases these proved to be substitution of every day language for estimation of quantitative amounts. The Ss varied greatly in quality and level of language and manifested a high degree of inconsistency in identifying quantitative amounts by verbal referents. In addition, the performance of the Ss on the remaining sections of the inventory indicate that a relatively small proportion of the students have capability of handling permutations and combinations, do not understand the concept of the independence of the probabilistic events and in general tend to avoid low probability events for prediction purposes. These data confirm the conclusions based on much of the prior research in this area and indicates a serious problem for both interpretation of "probability learning" studies as well as indicating a need for more systematic research on the educational programs and processes related to successful acquisition of these selected concepts of probability.

Study III: Estimation of Proportions and Correlation

The final study of this series was concerned with junior high school children's understanding of correlation and their ability to utilize information in estimating population parameters. The investigation of the latter was preceded by a more thorough analysis of information utilization and population estimation of college students (Belowicz, G., 1967) whereas the correlational study was based primarily upon earlier work by Piaget and his colleagues.

The tasks procedures were administered to individual Ss representing "high" and "low" achievers as defined by a relatively high performance on the Probability Inventory as well as indicies of intellectual ability. Table 15 presents the characteristics of the high and low Ss on each of the criterial variables.

Table 15

Mean Criterior Variable Scores for "High" and "Low" Subjects for Study III

	High N=12	Low N=12
Probability Inventory	23.2	13.8
Math Ach. Test	87%ile	67%ile
Math Grades	B	D+
IQ	115	99

The individual testing was completed in a small room at a local school over a period of three weeks. The testing sessions generally lasted about 50 minutes with all sessions recorded on tape. On entering the testing room the child was acquainted with the nature of the experiment and his attention was called to the tape recorder.

The testing session was divided into the following sections of Part I: Section 1, frequency estimates; Section 2a, predictions of a next "draw"; Section 2b, explanation by Ss of his estimations and predictions, and those that an imaginary person might make (E stated these); Section 3, inquiry to determine Ss basis and justification of estimations and a Section 4, the correlational task (Part II).

Part I: Estimation and the Information Utilization

The materials used in this part of the study are described fully in the dissertation mentioned earlier (Belovitz, G., 1967). Suffice to indicate here that each data display consisted of a 9 x 11½ sheet of paper on which were red and blue circles covered with clear plastic. The total of 24 displays were equally divided among the following types: a) proportion, 67-33 unordered (i.e., random placement), b) 67-33 proportions - ordered (i.e., alternating display), c) 50-50 unordered, and d) 50-50 ordered displays.

In each case the nature of the experiment was described briefly to the S and the hypothetical technique for sample selection was explained. S was asked to look at each display and tell the experimenter "about how many" red and blue elements he thought there would be in a total population of 1000 elements. S was then asked if he was to "draw" one element from the population of elements he was viewing whether he would get a red or blue element (Subjective belief). The experimenter than presented the S with the estimates and predictions of another (imaginary) person and asked the S why the person might make such estimates and predictions, finally a more full inquiry, varying with the nature of the S's answers, was instituted to determine S's basis or justification for his proportional estimates and subjective beliefs about the sampling process.

Results and Discussion. The results of this aspect of the study, which duplicates many of the essential forms

of that completed with college students (Belovitz, G., 1967) are as follows. First, there was no significant differences between the sophisticated ("high" group) and the unsophisticated Ss. While analysis of their verbal responses indicated a suggestive trend toward more accuracy (with respect to the process of estimation, etc.) the descriptive statistics indicated this was merely an impression gained from the more mature language of the sophisticated Ss. Second, in all essential respects the behavior of the junior high school Ss was not unlike that obtained with college students. The junior high school students report the same kinds of strategy, employ "Guessing" with a high degree of frequency, and generally give many indications of not comprehending the nature of the task or the principles involved.

These summary statements are confirmed by specific reference to several aspect of the data. First of all, there were a substantial number of incorrect proportion estimates (70%). The inquiry section yielded no systematic strategy accounting for these errors with the possibility, however that mere miscalculations, especially in the case of the 33-67 displays led to incorrect estimation. There were many cases where S was actually counting the sample elements, but not all Ss reported this as their procedure at arriving at a frequency estimation. During the early trials and/or in the absence of any other information (imaginary "draw") there was a definite preference for a 50-50 distribution of elements; there was an excess of 35% responses in this category. As indicated it was not possible to derive any indications of systematic differences in strategy between the sophisticated and unsophisticated groups. The inquiry yielded the following kinds of "strategies": 1) Sheer guess, 2) a "usual" proportion (generally 50-50), 3) reference to the proportions reported in the previous booklet, 4) preference for "even numbers", 5) color preference, 6) first number that came to mind.

Since the total number of booklets consisted of 12 with 50-50- and 12 with 33-67 distribution of elements it might be expected that Ss would move toward a stable level of frequency estimation and subjective belief. As one might expect the responses to the 50-50 order displays stabilized on trial 2 or 3. However, on 33-67 proportional displays only 1/2 of the Ss stabilized their proportional estimates prior to trial ten. It is quite probable that the computational difficulties mentioned above and the failure of the Ss to understand the nature of the task accounts for this. In many cases it was

quite obvious the Ss refused to "believe the displays" and would not respond to the actual proportional components of the display. The Ss did understand the relevance of 50-50 response in the absence of no further information, however. The data also indicated that Ss frequency estimation and predictions for the total set of displays did not closely match. This may have resulted from a strong tendency to derive each prediction from a particular display pattern and not accumulate the information across displays.

It seems warranted to suggest the following conclusions: it was clear that Ss ability to estimate population proportions was not very high. There were a substantial number of incorrect (60%) estimates most of which could be accounted for by the 33-67 display. Secondly, examination of the number of trials stabilization suggested that the Ss did not accumulate information across trial displays. Ss had a strong preference for 50-50 distribution or proportion which operated not only in the presence of objective information but also in the imaginary situation (absence of concrete information). Finally the subjective beliefs (predictions) of the Ss were derived from the immediately prior estimation but yielded no indication of probability matching when the responses to the set of booklets were considered.

Part II: Understanding of Correlation

The correlational task consisted of a set of specially constructed pictures of boys and girls which contained various combinations of hair and eye color as the criterial cues of association. The S was presented with the following 6 tasks:

(1) 12 pictures, 6 with brown hair and brown eyes and 6 with blue eyes and blonde hair

(2) 16 pictures, 4 with brown hair-brown eyes, 4 with brown hair-blue eyes, 4 with blonde hair-brown eyes and 4 with blonde hair-blue eyes.

(3) 12 pictures, 5 brown haired-brown eyes, 2 brown haired-blue eyes, 1 blonde hair-brown eyes, 4 blonde hair-blue eyes; (3a) S was asked to improve the "relationship".

(4) Two sets of 12 pictures were presented to the S. Set A consisted of 4 brown haired-brown eyed, 2 brown haired-blue eyed, 2 blue eyed-brown hair, 4 blue eyed-blonde hair, and Set B contained 3 brown hair-brown eyes, 3 brown hair-blue eyes, 1 blonde hair-brown eyes, 5 blonde haired-blue eyes.

(5) S was presented the total group of pictures (48) and asked to form two groups representing the same degree of relationship.

(6) S was given the total group of pictures and asked to form a group of 12 that indicated relationship among the physical features. S was asked to state the degree of "correlation" (task 1 and 2), to indicate the sameness or difference of two sets (task 4) or to improve the degree of relationship (3a) or to construct sets yielding a stated degree of relationship (tasks 5 & 6). In each case, the correctness of response and Ss justification or rational was recorded.

Results and Discussion. Table 16 presents the number of Ss in the High and Low group who responded correctly or incorrectly on each of the seven tasks. The table clearly indicates the high group achieved, overall, about twice as many correct and borderline responses as did the low Ss. Inspection of the data, task by task, shows the most difficult task was that requiring the formation of two groups with identical relationships among the criterial cues (task 6). The difficulty is evident from the performance of the Low group, and only 8 High Ss obtained "borderline" correct responses.

Analysis of the inquiry data yielded the major kinds of incorrect beliefs and/or response biases of the Ss that contributed to failure and interference with understanding of problem of correlation.

Task 1: The Ss stressed personal experiences and observations of the real world in responding to the initial task. For example, a most frequent response was that "most people with brown hair have brown eyes." Since this is a correct empirical generalization some of the initial correct responses on the task may have been given for incorrect reasons.

Table 16

Correct, Borderline and Incorrect Responses of Subjects in High and Low Groups on Correlational Task

	Sub-Tasks													
	1		2		3		3a		4		5		6	
	H	L	H	L	H	L	H	L	H	L	H	L	H	L
Correct	11	8	9	5	7	5	7	3	1	0	2	1	6	2
Borderline	0	1	1	1	3	2	3	3	5	1	8	3	3	2
Incorrect	1	3	2	6	2	5	2	6	6	11	2	8	3	8

Totals

	High	Low
Correct	43	24
Borderline	23	13
Incorrect	18	47

To check this possibility, a reversal of the contingency within the task groups of stimuli was presented (i.e., 6006 was changed to 0660 so that it was in contradiction to the child's bias). The reversal procedure reduced the number of correct responses by 25% in the high group and 50% in the low group.

About 30% of the children mentioned heredity as a reason for the relationship; several referred specifically to their general science class. The knowledge acquired in the classroom, of course, had the same effect as the subjective beliefs stated above. Finally, it was obvious that, with the exception of 5 Ss in the High group, the children had difficulty understanding the word "relationship". Much explanation and explication prior to the first task was necessary.

Task II: Ten of the children (6 high and 4 Low) perceived four groups (instead of one when this task was presented. That is, 4-4-4-4 set elicited a tendency to impose a structure on a sub-set and respond accordingly; e.g., "a strong relationship exists because all four children in one group has brown hair and brown eyes". The number of Ss involved prohibited any precise analysis, but other responses of these children suggests the grouping tendency might well be related to the child's understanding (or lack) of correlation.

Task III: Task III basically the same as task I and II - . S could complete sub-task I then one would expect he would do well on this one. Table 16indicates that III was slightly more difficult - only 7 in the high group and 5 in the low group obtained correct responses. No child obtained a correct response who had not obtained correct responses on Task I and II. The reason for incorrect responses by Ss who had responded correctly on I and II appeared to involve the attention to partial structure as in Task II; i.e., an increase in Ss uncertainty about the relationship because of different numbers of pictures in each criterial subgroup of the criterial attribute contingencies.

Task IIIa: In this task the child was asked to improve the relationship as presented in the immediately preceding task. In general, the children gradually built up to the ideal response (6006) adding or deleting cards one at a time. In several cases it was clear that Ss continued working at the task because E continued to ask questions (e.g., "can you make it even stronger?"). Thus, a number of children in each group achieved a borderline response that might not have done so otherwise. However, some of the children in the Low group, having strengthened the relationship, weakened it with their subsequent responses.

Task IV: The task comparing the set 4-2-2-4 with 3-3-1-5 proved to be one of the most difficult for these children. Only one child responded correctly and gave the correct reasons for his response. A major problem in this case was a strong preference for sub-group 5 (brown hair and brown eyes) which was consistent with Ss experentially derived biases.

Task V: This task proved too difficult for most of the children. E was able to get many of them to attempt the task and even to achieve borderline correctness (3 in the High group), but it was obvious this was a most demanding task requiring complete understanding of correlation.

Task VI: In task VI S was requested to form a group which had no relationship among the criterial cues. The children who were successful in this task used only two categories (for example, 1-1-0-0) and a minimum number of picture cards to form the group. Inquiry yielded indications the children had little or no understanding of correlation.

These data are based on relatively few number of Ss and conclusions from them would be hazardous to say the least. At the same time, the differentiation of High and Low Ss on this task follows that obtained on the Probability Inventory. It is therefore reasonable to suspect the tasks require similar levels of cognitive development to achieve adequate performance.

Further, the data confirm Piaget's prior work on children's understanding of correlation - junior high school children have not yet obtained an adequate understanding of correlation. The extensive inquiry data obtained yielded many clues as to the reasons for Ss' difficulties with the task that need to be followed by more systematic studies.

SUMMARY AND CONCLUSIONS

The introduction of the new mathematics in elementary and junior high education during recent years has given impetus to many new areas of investigation relevant to educational psychology. The evaluation of its impact on learning and cognitive growth is both complex and controversial. A major question is what experience is necessary (stage of cognitive development) for the child to have the requisite cognitive skills for understanding the quantitative relationships to be learned. The axioms of probability theory have counter-parts in the concepts of set theory and therefore what one learns about the understanding of probability is directly applicable to the understanding of the set theoretic concepts which form the basis for most of the new mathematics programs. Whether or not younger children have capacity for understanding probability has become highly controversial. Many statements in experimental literature imply such understanding, whereas the work of Piaget and others indicate children's understanding of probability concepts is highly questionable. The purpose of this study was to investigate junior high school children's understanding of the concept intrinsic to the theory of probability. The specific purposes were to determine:

1) the effects of variation and consistency in usage at quantitative language terms by junior high school children.

2) children's ability to generate potential combinations and permutations and recognition of all possible outcomes of a particular set of actions.

3) children's degree of recognition and utilization of the independence of probabilistic events.

4) the extent to which a preference for low probability events determined responses.

5) children's ability to appropriately assess and modify contingency relationships.

6) the effect of variation in event structures of probabilistic events on estimation of population parameters and subjective certainty.

To achieve these objectives a series of specially constructed tasks were used. The various tasks were administered to a group of 36 college and 35 junior high school students. On the basis of preliminary study, the investigation was divided into three parts:

1) A study of the effect of variation in event structure on estimation of population parameters and subjective certainty (Objective 6);

2) an inventory of concepts of probability was constructed and administered to junior high school students to study the following aspects of probability theory: a) variation and consistency in usage of quantitative language terms (Objective 1); b) ability to generate combinations and permutations (Objective 2); c) recognition and utilization of the concept of independence of events (Objective 3); and d) the effects of low probability events on decision outcomes (Objective 4).

3) An individually administered task was designed to determine: a) junior high school children's ability to estimate population parameters under variation in the stimulus set (Objective 6), and b) to assess and modify contingency relationships (Objective 5).

The results of Study 1 indicated: 1) college students were able to discriminate between 50-50 and 33-67 population proportions; accuracy was highest for the 50-50 alternating probabilistic displays. 2) College students tend to over estimate proportions above .50 and underestimate proportions below .50. 3) Large sample sizes were associated with decrease subjective certainty as were randomization of display elements. Six major strategies were used by college students in the estimation of population proportions. The choice of strategy appeared to depend on the Ss "desire for certainty", and subjective beliefs about the estimation of proportions. It was concluded (see Belowicz, G., 1967) that responses to the probabilistic situations could not be taken at face value; i.e., labeled "good" or "bad", "appropriate" or "inappropriate" on the basis of the obtained quantitative indices of performance. The rationality of the

response could only be evaluated by investigating the underlying decision processes and the subjective beliefs of the Ss through intensive inquiry. In many cases it was clear that "correct responses" and predictions were derived from erronious subjective beliefs. Ss did not evidence a high level of sophistication or understanding of the requisite probability concepts. These findings contradict the assumptions made by many investigators concerned with probability behavior and learning. The data was consistent with the findings of Piaget and others who have systematically and directly investigated children's and adult's concepts of probability.

The data obtained from this study with college students was used to construct similar tasks for obtaining similar data with junior high school children. Such data was obtained from a group of "high" and "low" achievers (Study 3, Part 1). The younger children used strategies in estimating population proportions and held correct and erronious subjective beliefs about probabilistic events similar to those of college students.

The second sub-study (Study 2) utilized a specially constructed inventory to survey characteristic usage of quantitative language terms, understanding of permutations and combinations, understanding of independence of probabilistic events and use of low probability events of 63 junior high school children. The results indicated that intelligence level, mathematics achievement, and sex of subject were significant determinants of performance level. The children had a wide variety of available quantitative terms but over 50% of these terms were neither precise nor used with a high degree of consistency. A relatively small proportion of the students were capable of handling permutations and combinations, few understood the concept of independence and generally did not utilize low probability events appropriately for prediction purposes. These data offer general confirmation to the conclusions based on prior research and indicate a serious problem for interpretation of many probability learning studies as well as a need for more systematic research on processes related to successful acquisition of these selected concepts of probability. A study currently underway indicates, for example, that not until 16 or 17 do as many as 30% of high school children systematically utilize one of the several rules or strategies for efficient generation of permutations.

The third study in this investigation compared children who had High ($N = 12$) and Low ($N = 12$) performance

on the Probability Inventory as well as differentiable on the basis of IQ and mathematics grades. As indicated earlier the criterial variables did not differentiate the children on their capability of estimating proportions nor did it yield any evidence of more or less understanding of probability estimation and then that obtained with the college students. The second part of Study Three was designed to determine the ability of the children in the High and Low groups to modify contingency relationships. Results indicated that junior high school children have an inadequate understanding of the concept of correlation. In most cases the Ss were unable to deal with the contingency of events separate from rules generated from their individual or educational experiences (e.g., that brown hair and brown eyes go together because of heredity).

Taken together, the findings of the three studies indicate a need for more systematic investigation of specific concepts underlying probability. In each of the areas studied it seems clear that both college and junior high school students, who presumably have sufficient educational experience to understand probability concepts, have not acquired the rules or strategies of responding to probabilistic situations that correspond to the concepts underlying probability. It would appear that "probabilistic learning", and even mathematics learning, does not have the wide generality among young college students and junior high school students that might be expected on the basis of their educational experiences. Further intensive study of the processes underlying the concepts of probability should yield a better understanding of the cognitive processes involved as well as provide the basis for the development of educational procedures and curricula that provide more lasting learning and transfer of these concepts to new situations.

REFERENCES

Annett, M. The classification of instances of four common class concepts by children and adults. Brit. J. educ. Psych., 1959, 29, 223-236.

Bennet, W. F., Fitts, P. M., & Merrill, N. The learning of sequential dependencies. J. exp. Psych., 1954, 48, 303-312.

Brim, O. G. & Koenig, F. W. Two aspects of subjective probability among college students. J. Communications, 1959, 9, 19-26.

Brunk, H. D. An Introduction to Mathematical Statistics. Boston: Gill & Co., 1960.

Brunswik, E. Probability as a determiner of rat behavior. J. exp. Psych., 1939, 25, 175-197.

Brunswik, E. Organismic achievement and environmental probability. Psych. Rev., 1943, 50, 255-272.

Cohen, J. & Hansel, M. Risk and Gambling. New York: Philosophical Library Inc., 1956.

Craig, Grace J. & Myers, J. L. A developmental study of sequential two-choice decision making. Child Developm., 1963, 34, 483-493.

Dale, H. C. A. On choosing between bets. Quart. J. exp. Psych., 1962, 14, 49-51.

Edwards, W. Probability learning in 1000 trials. J. exp. Psych., 1961, 62, 385-394.

Edwards, W. A bibliography of decision making. Unpublished, 1964.

Edwards, W. & Phillips, L. D. Man as a transducer for probabilities in Bayseian command and control systems. In G. L. Bryan & M. W. Shelly (Eds.), Human Judgments and Optimality. New York: Wiley, 1964.

Estes, W. K. & Straughan, J. H. Analysis of a verbal conditioning situation in terms of statistical learning theory. J. exp. Psych., 1954, 47, 225-234.

Gratch, G. The development of the expectation of the non-independence of random events in children. Child Developm., 1959, 30, 217-227.

Inhelder, B. & Piaget, J. The Growth of Logical Thinking from Childhood to Adolescence. Basic Books, 1958.

Jenkins, W. O. & Stanley, J. S. Partial reinforcement: a review and critique. Psych. Rev., 1950, 47, 193-234.

Jones, Mary H. & Liverant, S. Effects of age differences on choice behavior. Child Developm., 1960, 31, 673-680.

Kass, N. Risk in decision making as a function of age, sex, and probability preference. Child Developm., 1964, 35, 577-582.

Messick, S. J. & Solley, C. M. Probability learning in children: some exploratory studies. J. genet. Psych., 1957, 90, 23-32.

Middleton, Jane. Some personality correlates of decision processes in children -- a convergent and predictive validity study. Unpublished Ph.D. Dissertation, Purdue University, 1963.

Pratt, K. C., Hartman, W. E. & Mead, J. L. Indeterminate number concepts. III. Representation by children through selection of appropriate aggregations. J. genet. Psych., 1954, 84, 39-63.

Rubin, B. A developmental analysis of decision making behavior. Unpublished Ph.D. Dissertation, Purdue University, 1964.

Shantz, D. Probabilistic classification learning. Unpublished Ph.D. Dissertation, Purdue University, 1965.

Smedlund, J. Concept of correlation in adults. Scand. J. Psych., 1963, 4, 165-173.

Smock, C. D. Effects of Motivational Factors on Perceptual-Cognitive Efficiency of Children Who Vary in Intellectual Level. Cooperative Research Project No. 790.

Stevenson, H. W. & Odom, R. D. Children's behavior in a probabilistic situation. J. exp. Psych., 1964, 68, 260-268.

Stevenson, H. W. & Zigler, E. F. Probability learning in children. J. exp. Psych., 1958, 56, 185-192.

Tolman, E. C. & Brunswik, E. The organism and the causal texture of the environment. Psych. Rev., 1935, 42, 43-77.

Uhl, C. N. Two choice probability learning in the rat as a function of incentives, probability of reinforcement, and training procedures. J. exp. Psych., 1943, 62, 63-76.

Welch, L. & Long, L. Comparison of the reasoning ability of two age groups. J. genet. Psych., 1943, 62, 63-76.

Yost, Patricia A., Siegel, Alberta E., & Andrews, Julia M. Nonverbal probability judgments by young children. Child Developm., 1962, 33, 769-780.

APPENDIX A

Definitions and Axioms Underlying Probability

The elements of a probability space are called elementary events. Certain collections of sets of elementary events are events. The family T of all events has the following properties:

T1. O and S are events; O is the void event, containing no elementary events, and S is the sure event, containing all elementary events.

T2. The intersection of a countable set of events is an event.

T3. The union of a countable set of events is an event.

T4. The complement, $\bar{A}$, of an event A is an event.

T5. The difference of two events is an event.

If there are only a finite number of elementary events, then every set of elementary events is an event, and the axioms for T are satisfied.

Associated with each event A is a probability P(A) with the following properties.

P1. P(A) is defined for every event A.

P2. $P(A) \geq 0$ for every event A.

P3. $P(S) = 1$.

P4. The probability of the union of a countable set of mutually exclusive events is the sum of their probabilities.

P5. $P(B - A) = P(B)$ for every pair A,B of events for which $A < B$.

P6. $\underline{P}(\bar{\underline{A}}) = 1 - (\underline{P}(\underline{A})$ for every event $\underline{A}$.

P7. $\underline{P}(\underline{O}) = O$.

P8. $O \leq P(A) \leq 1$ for every event $\underline{A}$.

P9. $\underline{P}(\underline{A}) \leq \underline{P}(\underline{B})$ if $\underline{A} < \underline{B}$.

The <u>conditional</u> <u>probability</u> $\underline{P}(\underline{B}/\underline{A})$ is defined by:

$$\underline{P}(\underline{B}/\underline{A}) = \underline{P}(\underline{AB})/\underline{P}(\underline{A}).$$

Two events $\underline{A}$ and $\underline{B}$ are <u>independent</u> if

$$\underline{P}(\underline{AB}) = \underline{P}(\underline{A})\underline{P}(\underline{B}).$$

The events of a countable collection of events are independent if the probability of the intersection of any finite number of them is the product of their probabilities.

Bayes formula applies to a situation $A < \overset{N}{\underset{i}{}} = {}_1 B$,

and where the events $B_1, B_2, \ldots, B_n$ are mutually exclusive.

Then we have for each k, k = 1, a, . . . , N.

$$P(B_k/A) = \frac{P(B_k)P(A/B_k)}{\sum_{i=1}^{N} P(B_1)P(A/B_1)}$$

Brunk, H. D. <u>An Introduction to Mathematical Statistics</u>, Boston: Ginn and Company, 1960, p. 31.

APPENDIX B

Table B1

Descriptive Data on College and Junior High School Students

Age	Sex	
	Male	Female
10 - 12	1	0
13 - 15	30	28
16 - 18	4	0
19 - 21	22	9
22 - 24	4	1

APPENDIX C

Language Responses to Fractions, Percentages and Absolute Numbers in Part I of Probability Inventory

Low 1.	almost none couple fewer, fewest hardly any pretty small shortest	hardly many hardly very much less than the smallest little, little bit smallest, smaller tiny(bit)	minute none not hardly any not very big(at all) scarcely very few, little	not many at all not so very many not too much point sliver etc.
Level 2.	almost half less low not too small small	bigger than small less than half(part) not much next to smallest small amount	few little group not many part of a part small but not tiny	fairly small little less than ½ not too big short small part
Level 3.	almost good sized between half(for absolutes only) not big but not small wee more than ½	about ½ cake cut down middle medium one side of a pie some	almost medium equal(amt.) middle(sized) piece some	average(amt.) group just about right part

APPENDIX C (cont'd)

Level 4.	big	bigger than between(½)	considerably large	part
	enough	fairly large	fat	good sized
	greater than	good amount	gob	high
	heavier	large	little(bit)more	little bigger
	large group	majority	many	more
	more than part (½)	next to largest	not all	pretty much
	pretty good	over	quite a few	several
	that much	tall	thick	
High 5.	about all	all	almost(all)	almost full
	almost whole	a lot	awful lot	bigger, biggest
	giant	huge	highest	he-man
	fairly bigger	less than 1		less than all
	longer	larger, largest	lots	more than large
	more than enough	mamoth	many more	most
	just about all	not quite all	real big, large, etc.	scores
	tallest	very large pie	very large, big, etc.	
	whole bunch			

OE 6000 (REV. 9-66)

DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
OFFICE OF EDUCATION

ERIC REPORT RESUME

ERIC ACCESSION NO.

CLEARINGHOUSE ACCESSION NUMBER

RESUME DATE: 02-1-68 | P.A. | T.A.

IS DOCUMENT COPYRIGHTED? YES ☐ NO ☐

ERIC REPRODUCTION RELEASE? YES ☒ NO ☐

TITLE

UNDERSTANDING OF CONCEPTS OF PROBABILITY THEORY BY JUNIOR HIGH-SCHOOL CHILDREN

PERSONAL AUTHOR(S)

Smock, Charles D. Belovicz, Gretchen.

INSTITUTION (SOURCE)

Purdue Univ., Lafayette, Ind. Dept. Child Development

SOURCE CODE

REPORT/SERIES NO.

OTHER SOURCE

SOURCE CODE

OTHER REPORT NO.

OTHER SOURCE

SOURCE CODE

OTHER REPORT NO.

PUB'L. DATE 28-Feb-68 | CONTRACT/GRANT NUMBER OEG-3-6-06125-1739

PAGINATION, ETC.

60p.

RETRIEVAL TERMS

Probability theory, Children's learning, Cognitive development

IDENTIFIERS

ABSTRACT

A major question of concern to the "new mathematic" educators is what cognitive skills are necessary for understanding the quantitative concepts and relationships to be learned. The quantitative relationships involved in probability theory can be subsumed under "set theory" which is also the "content" of the "new" mathematics. What one learns about the understanding of probability is applicable, therefore, to the understanding of set theoretic concepts.

The primary purpose of this study was to investigate junior-high school children's understanding of the concepts intrinsic to the theory of probability. The specific objectives were to explore: 1) Variation and consistency in usage of quantitative language terms; 2) Ability to generate possible combinations and permutations; 3) Recognition and utilization of the concept of independence of events; 4) Ability to appropriately assess, and modify, contingency relations in a complex set of related events; and 5) Effect of variation in event-structures of probabilistic event-sets on estimation of population parameters and subjective certainty of these estimations.

Experimental procedures were designed to obtain data on each of these objectives. Results indicated that children (even college students) fail to understand the basic ideas of probability theory and that intensive study of particular concepts (e.g., permutations, correlation) and the experience necessary for acquisition and a utilization of appropriate quantitative language terms should be undertaken